

Partially observed Markov random fields are variable neighborhood random fields

M. Cassandro

A. Galves

E. Löcherbach

April 17, 2012

Abstract

The present paper has two goals. First to present a natural example of a new class of random fields which are the variable neighborhood random fields. The example we consider is a partially observed nearest neighbor binary Markov random field. The second goal is to establish sufficient conditions ensuring that the variable neighborhoods are almost surely finite. We discuss the relationship between the almost sure finiteness of the interaction neighborhoods and the presence/absence of phase transition of the underlying Markov random field. In the case where the underlying random field has no phase transition we show that the finiteness of neighborhoods depends on a specific relation between the noise level and the minimum values of the one-point specification of the Markov random field. The case in which there is phase transition is addressed in the frame of the ferromagnetic Ising model. We prove that the existence of infinite interaction neighborhoods depends on the phase.

Key words : Random lattice fields, variable neighborhood random fields, Ising model.

AMS Classification : Primary: 60G60, 60K35 Secondary: 82B20, 82B99

1 Introduction

Recent experimental data suggest that populations of neurons have interactions of variable range. There are reasons to believe that the interaction neighborhood of each neuron is not fixed, but changes as a function of the configuration. Actually the same phenomenon seems to be present at different scales when instead of individual neurons local sub-populations of neurons act as interaction unity. Justifying the variable neighborhood assumption for fields describing populations of neurons is an important open question in neuroscience. For a general discussion of the geometry of the neuronal connectivity we refer the reader to Braitenberg and Schütz (1998). Concerning the relationship between the time evolution of the neuronal activity and the reaction to external stimulations see MacLean et al. (2005). For a very recent statistical and clinical discussion of the way neighborhood interactions between regions of the brain can change we refer to Wang et al. (2010). Finally, for a mathematical model describing variable range interactions in time rather than in space we refer to Cessac (2011) and the references cited therein.

The above observation suggests to model these kind of interacting systems by a new class of random fields which are the variable neighborhood random fields. This new class of

models is a natural extension to the case of random fields of the notion of stochastic chains with memory of variable length introduced by Rissanen (1983).

Random fields with variable interaction neighborhoods have recently gained interest, and some papers are devoted to the study of such kind of new models, see Dereudre et al. (2011) and Löcherbach and Orlandi (2011). The first paper focusses on the problem of existence of these models in \mathbb{R}^d . The second paper addresses the problem of statistical inference, mainly in the case of bounded interaction range.

The present paper has two goals. First we present a simple and natural class of variable neighborhood random fields, namely the incompletely observed Markov random fields. The second goal is to search for sufficient conditions ensuring that the variable neighborhoods are simultaneously finite for almost every realization of the field.

The model we consider is a nearest neighbor Markov random field taking the values $+1$ or -1 . At each site there is an independent random mechanism which hides the actual value of the spin and replaces it in the observed data by the value -1 . This can be seen as a black and white picture in which random noise affects the readability of some of the pixels which appear black independently of the actual color. In particular, the noise mechanism introduces a high bias into the system. As in the one-dimensional case where random observations of Markov chains lead to processes having infinite memory, see e.g. Collet and Leonardi (2009), a priori such a model is a random field having infinite memory. However, in this particular case, the partially observed Markov random field is indeed a variable neighborhood random field, and the relevant neighborhoods needed in order to determine the spin at a given site will be regions surrounded by a circuit of sites having all spins equal to $+1$. This is the content of Theorem 1.

Several questions arise naturally in this context. First, is there a relation between presence or absence of phase transition for the underlying random field in \mathbb{Z}^2 and finiteness of the interaction regions of the variable neighborhood random field? Does the absence of phase transition always imply that the interaction regions are finite almost surely? Do infinite interaction regions always exist in the regime of phase transition? It turns out that the question of presence/absence of phase transition and the question of finiteness of interaction neighborhoods are related in a more intricate way than we would have guessed naively.

The case in which there is no phase transition is treated in our Theorem 2. If the minimum values of the one-point specification of the original Markov random field are large enough, we show that the two situations are possible, depending on the specific relationship between the perturbation level and the specification minima. The proof relies on a coupling argument.

The case in which there is phase transition is addressed in the frame of partially observed ferromagnetic Ising models. We show that, as a consequence of the bias of the noise, the plus phase and the minus phase behave differently when the perturbation level and the temperature are small enough. Namely, in the plus phase all interaction neighborhoods will be finite almost surely, while in the minus phase, infinite interaction regions will always exist with strictly positive probability. This is the content of Theorem 3. The proof of this theorem is based on a Peierls contour counting argument.

This paper is organized as follows. Definitions, notation and main results are presented in Section 2. The proofs of Theorem 1, Theorem 2 and Theorem 3 are presented successively in Sections 3, 4 and 5.

2 Definitions, notation and main results

Let $A := \{-1, 1\}$ and $S = A^{\mathbb{Z}^2}$ be the set of all possible configurations. We endow S with the product sigma algebra \mathcal{S} . Fixed configurations will be denoted by lowercase letters x, y, z . A point $i \in \mathbb{Z}^2$ is called a site.

If $x \in S$ is a configuration, then for any $i \in \mathbb{Z}^2$, x_i will denote the value of the configuration at site i . Given a subset $F \subset \mathbb{Z}^2$, we will also denote $x_F = \{x_i, i \in F\}$. Let $X = \{X_i : i \in \mathbb{Z}^2\}$ be the collection of projections on S , defined by $X_i(x) = x_i$ for all $i \in \mathbb{Z}^2$. We introduce the following σ -algebras: For any $\Lambda \subset \mathbb{Z}^2$, let

$$\mathcal{F}_\Lambda = \sigma\{X_i : i \in \Lambda\}.$$

Definition 1 *Any probability measure on (S, \mathcal{S}) will be called a random field.*

We consider random fields on (S, \mathcal{S}) which are defined by their local specifications, see Dobrushin (1970). In order to do so, we recall the notion of specification from Georgii (1988).

Definition 2 *A specification on (S, \mathcal{S}) is a family $P = \{P_\Lambda\}_{\Lambda \subset \mathbb{Z}^2}$ of probability kernels on (S, \mathcal{S}) such that*

- (a) *For each $\Lambda \subset \mathbb{Z}^2$ finite and each $B \in \mathcal{S}$, the function $P_\Lambda(B \mid \cdot)$ is \mathcal{F}_{Λ^c} -measurable.*
- (b) *For each $\Lambda \subset \mathbb{Z}^2$ finite and each $B \in \mathcal{F}_{\Lambda^c}$, $P_\Lambda(B \mid y) = 1_B(y)$.*
- (c) *For any pair of finite subsets Λ and Δ , with $\Lambda \subset \Delta \subset \mathbb{Z}^2$, and any measurable set B ,*

$$\int P_\Lambda(B \mid z) P_\Delta(dz \mid y) = P_\Delta(B \mid y) \quad (2.1)$$

for all $y \in S$.

From (a) and (b) above it follows that P_Λ can be identified with probability weights $p_\Lambda(x_\Lambda \mid y_{\Lambda^c})$ such that for any $B \in \mathcal{F}_\Lambda$,

$$P_\Lambda(B \mid y) = \sum_{x_\Lambda : (x_\Lambda, y_{\Lambda^c}) \in B} p_\Lambda(x_\Lambda \mid y_{\Lambda^c}).$$

Definition 3 *A random field μ on (S, \mathcal{S}) is consistent with a specification P if for each finite subset $\Lambda \subset \mathbb{Z}^2$,*

$$\int \mu(dx) P_\Lambda(B \mid x) = \mu(B), \quad (2.2)$$

for every $B \in \mathcal{S}$. We write $\mathcal{G}(P)$ for the set of all random fields consistent with the specification P .

In the sequel, if P is a specification, for $\Lambda = \{i\}$, instead of writing $p_{\{i\}}(\cdot \mid \cdot)$ for the probability weights, we shall use the short-hand notation $p_i(\cdot \mid \cdot)$.

One important class of random fields are the Markov fields. For the reader's convenience we recall here the basic definitions. The L^1 -norm $\|\cdot\|_1$ is defined as follows. If $i = (i_1, i_2) \in \mathbb{Z}^2$, then

$$\|i\|_1 = |i_1| + |i_2|.$$

For any subset $F \subset \mathbb{Z}^2$, the boundary ∂F is defined as

$$\partial F = \{j \in F : \exists k \in \mathbb{Z}^2 \setminus F : \|j - k\|_1 = 1\}.$$

Definition 4 *Two points i and $j \in \mathbb{Z}^2$ are called L^1 -neighbors if $\|i - j\|_1 = 1$.*

Definition 5 *Let P be a specification. $\mu \in \mathcal{G}(P)$ is a Markov random field of order 1 if for all $i \in \mathbb{Z}^2$, the function $P_i(\{X_i = +1\}|\cdot)$ is $\mathcal{F}_{\partial(\{i\}^c)}$ -measurable, where $\partial(\{i\}^c) = \{j \in \mathbb{Z}^2 : \|i - j\|_1 = 1\}$.*

We now introduce the notion of variable neighborhood random fields.

Definition 6 *Let P be a specification and $\mathbb{P} \in \mathcal{G}(P)$. We say that \mathbb{P} is a variable neighborhood random field if for any $i \in \mathbb{Z}^2$ there exists a mapping $C_i : A^{\mathbb{Z}^2 \setminus \{i\}} \rightarrow \mathcal{P}(\mathbb{Z}^2 \setminus \{i\})$ such that the following statements hold.*

1. *For any site $i \in \mathbb{Z}^2$ and any subset $\Lambda \subset \mathbb{Z}^2$, the event $\{C_i(X) = \Lambda\}$ belongs to \mathcal{F}_Λ .*
2. *For all $i \in \mathbb{Z}^2$, the map $x \mapsto P_i(\{+1\}|x)$ is $\mathcal{F}_{C_i(X)}$ -measurable. Here,*

$$\mathcal{F}_{C_i(X)} = \{B \in \mathcal{S} : B \cap \{C_i(X) = \Lambda\} \in \mathcal{F}_\Lambda, \text{ for all } \Lambda \subset \mathbb{Z}^2\}.$$

3. *$C_i(x)$ cannot be shortened. This means that if there is another collection of maps $\tilde{C}_i, i \in \mathbb{Z}^2$, such that the above properties hold, then $C_i(x) \subset \tilde{C}_i(x)$ for all i and x .*

We call $x_{C_i(x)}$ the context of site i , given the configuration x . The functions $\{C_i(\cdot), i \in \mathbb{Z}^2\}$ will be called context support functions of the variable neighborhood random field.

From now on we shall write shortly VNRF for variable neighborhood random field.

The goal of this paper is to introduce a natural example of VNRF's, namely incompletely observed Markov random fields. By this we mean the following. For any fixed $\varepsilon \in]0, 1[$, let ν_ε be the Bernoulli product measure

$$\nu_\varepsilon = \prod_{i \in \mathbb{Z}^2} (\varepsilon \delta_{-1} + (1 - \varepsilon) \delta_{+1}) \quad (2.3)$$

on (S, \mathcal{S}) . Thus under ν_ε , the coordinates $X_i, i \in \mathbb{Z}^2$, are i.i.d. random variables taking the value $+1$ with probability $1 - \varepsilon$ and the value -1 with probability ε . For each site i , its original color chosen according to μ will be observed only with probability $1 - \varepsilon$, and with probability ε , we loose any information concerning the color and report as output the value -1 .

Mathematically speaking, this means the following. For any measure $\mu \in \mathcal{G}(P)$ we consider the product measure $\mu \otimes \nu_\varepsilon$ on $(S \times S, \mathcal{S} \otimes \mathcal{S})$ and consider the probability measure \mathbb{P}^ε on

(S, \mathcal{S}) which is the image measure of $\mu \otimes \nu_\varepsilon$ under the operation of taking the point-wise minimum

$$S \times S \ni (x^1, x^2) \mapsto x \in S : \text{ for all } i \in \mathbb{Z}^2, x_i = x_i^1 \wedge x_i^2. \quad (2.4)$$

In other words, observing a realization of the random field \mathbb{P}^ε amounts to saying that we observe a realization of the original random field μ , where for each site, independently of the original value of the field and independently of the other sites, its value is replaced by the value -1 .

It turns out that incompletely observed random fields as defined above are VNRF's. This is the content of our first theorem. Before stating it, we recall some definitions. For any finite subset $F \subset \mathbb{Z}^2$, the interior \mathring{F} is given by

$$\mathring{F} = F \setminus \partial F.$$

Definition 7 *A self-avoiding path in \mathbb{Z}^2 is a finite sequence $\Gamma = (i_1, \dots, i_n)$ of sites such that i_j and i_k are L^1 -neighbors if and only if $|j - k| = 1$.*

Recall also that a set $F \subset \mathbb{Z}^2$ is called L^1 -connected if for any pair of points $i, i' \in F$, $i \neq i'$, there exists an integer $n \geq 1$ and a self-avoiding path (i_1, \dots, i_n) of sites in F such that $i_1 = i$ and $i_n = i'$.

The following theorem holds.

Theorem 1 *If $\mu \in \mathcal{G}(P)$ is a Markov random field of order 1 and ν_ε the Bernoulli probability measure of (2.3), then the measure \mathbb{P}^ε defined in (2.4) is a VNRF, and the context support functions are given by*

$$C_i(x) = \left(\bigcap \{F \subset \mathbb{Z}^2 : i \in \mathring{F}, F \text{ is } L^1\text{-connected and } x_{\partial(F)} \equiv +1\} \right) \setminus \{i\},$$

if there exists at least a finite set satisfying these conditions. Otherwise, we put $C_i(x) = \mathbb{Z}^2 \setminus \{i\}$.

It is natural to ask whether for a given Markov random field model μ and a given ε , all contexts $C_i(x)$ will be finite almost surely or not. In the regime of absence of phase transition, a first answer can be given by using coupling arguments. Call

$$\lambda_0^+ = \inf_{i \in \mathbb{Z}^2} \inf_{x \in S} p_i(+1|x), \quad \lambda_0^- = \inf_{i \in \mathbb{Z}^2} \inf_{x \in S} p_i(-1|x). \quad (2.5)$$

Let p^* be the critical probability for the site percolation model in \mathbb{Z}^2 . For a general presentation of percolation models we refer the reader to the classical treatise by Grimmett (1999). Then we have the following result.

Theorem 2 *Grant the conditions of Theorem 1. If*

$$(1 - \varepsilon)\lambda_0^+ > 1 - p^*, \quad (2.6)$$

then

$$\mathbb{P}^\varepsilon \left(\bigcap_{i \in \mathbb{Z}^2} \{|C_i(X)| < \infty\} \right) = 1.$$

On the other hand, if

$$\varepsilon + (1 - \varepsilon)\lambda_0^- > p^*, \quad (2.7)$$

then

$$\mathbb{P}^\varepsilon \left(\bigcup_{i \in \mathbb{Z}^2} \{|C_i(X)| = \infty\} \right) > 0.$$

In the above statements, $|C_i(X)|$ means the cardinal of the set $C_i(X)$. We give the proof of this theorem in Section 5 below.

We now consider the regime in which there is phase transition. We address the question of finiteness of contexts in the framework of the ferromagnetic Ising model.

Definition 8 *The homogeneous ferromagnetic Ising model is defined by the following specification. For any $\beta \geq 0$, $x, y \in S$ and any finite subset $\Lambda \subset \mathbb{Z}^2$,*

$$p_{\Lambda, \beta}(x_\Lambda | y_{\Lambda^c}) = \frac{1}{Z_{\Lambda, \beta}^y} \exp \left(\beta \left[\frac{1}{2} \sum_{i \in \Lambda} \sum_{j \in \Lambda, \|i-j\|_1=1} x_i x_j + \sum_{i \in \Lambda} \sum_{j \in \Lambda^c, \|i-j\|_1=1} x_i y_j \right] \right), \quad (2.8)$$

where

$$Z_{\Lambda, \beta}^y = \sum_{x_\Lambda \in A^\Lambda} \exp \left(\beta \left[\frac{1}{2} \sum_{i \in \Lambda} \sum_{j \in \Lambda, \|i-j\|_1=1} x_i x_j + \sum_{i \in \Lambda} \sum_{j \in \Lambda^c, \|i-j\|_1=1} x_i y_j \right] \right).$$

It is well known, see for instance Georgii (1988) and Presutti (2009), that there exists a critical value β^c , such that for all $\beta > \beta^c$ the set $\mathcal{G}(P)$ contains two extremal measures μ_β^- and μ_β^+ which are the pure states obtained by passing to the limit $\Lambda \rightarrow \mathbb{Z}^2$, taking the external configuration $y_j = -1$ for all $j \in \Lambda^c$ and $y_j = +1$ for all $j \in \Lambda^c$, respectively.

We write $\mathbb{P}_{\beta, \varepsilon}^+$ for the image measure of $\mu_\beta^+ \otimes \nu_\varepsilon$ under the map

$$S \times S \ni (x^1, x^2) \mapsto x \in S : \text{ for all } i \in \mathbb{Z}^2, x_i = x_i^1 \wedge x_i^2.$$

$\mathbb{P}_{\beta, \varepsilon}^-$ is defined in an analogous way.

We have the following result.

Theorem 3 *The following two statements hold.*

1. *For sufficiently large $\beta > \frac{1}{2} \ln 3$ and $\varepsilon < \frac{1}{3} - e^{-2\beta}$,*

$$\mathbb{P}_{\beta, \varepsilon}^+ \left(\bigcap_{i \in \mathbb{Z}^2} \{|C_i(X)| < \infty\} \right) = 1.$$

2. *For all $\beta > \beta^c$,*

$$\mathbb{P}_{\beta, \varepsilon}^- \left(\bigcup_{i \in \mathbb{Z}^2} \{|C_i(X)| = \infty\} \right) = 1.$$

Remark 1 *The high-temperature results of Theorem 2 apply very nicely in the framework of the ferromagnetic Ising model. In this case,*

$$\lambda_0^+ = \lambda_0^- = (1 + \exp(8\beta))^{-1},$$

and for example condition (2.6) reads as

$$\varepsilon < p^* \text{ and } \beta < \frac{1}{8} \ln \left(\frac{1 - \varepsilon}{1 - p^*} - 1 \right).$$

3 Proof of Theorem 1

Write P^ε for the specification of \mathbb{P}^ε and p_i^ε for the associated one-point conditional probabilities. The strategy of our proof is to provide an explicit formula for $p_i^\varepsilon(+1|x)$ (see (3.9) below) which shows that $p_i^\varepsilon(+1|x)$ depends only on $x_{C_i(x)}$.

Let $\Lambda_n = [-n, n]^2$. Write $\bar{C}_i(X) = C_i(X) \cup \{i\}$. Let $B \in \mathcal{S}$. We only have to consider the event $\{|C_i(X)| < \infty\}$. We start by evaluating

$$\mathbb{P}^\varepsilon(B; |C_i(X)| < \infty) = \lim_n \mathbb{P}^\varepsilon(B; C_i(X) \subset \Lambda_n).$$

We have for any fixed n ,

$$\mathbb{P}^\varepsilon(B; C_i(X) \subset \Lambda_n) = \sum_{F \subset \Lambda_n} \mathbb{P}^\varepsilon(B; \bar{C}_i(X) = F).$$

Here, we sum over all sets $F \subset \Lambda_n$ which are L^1 -connected and contain i in their interior. We evaluate each of these terms $\mathbb{P}^\varepsilon(B; \bar{C}_i(X) = F)$, for a fixed set F . In order to simplify notation, we write $\Lambda = \overset{\circ}{F}$ for the interior of F . Notice that $i \in \Lambda$.

Observe that

$$\mathbb{P}^\varepsilon(B; \bar{C}_i(X) = F) = \sum_{z_F \in A^F: \bar{C}_i(z) = F} \mathbb{P}^\varepsilon(B, X_F = z_F).$$

For any fixed F and z_F , notice that we can rewrite $1_B(x)1_{z_F}(x_F) = 1_{B^{z_F}}(x)1_{z_F}(x_F)$, where

$$B^{z_F} = \{x \in S : (z_F, x_{F^c}) \in B\} \in \mathcal{F}_{F^c}.$$

Hence, using Fubini's theorem and since $z_{\partial F} \equiv +1$, by construction of \mathbb{P}^ε ,

$$\begin{aligned} \mathbb{P}^\varepsilon(B; X_F = z_F) &= \int_S \nu_\varepsilon(dx^1) 1_{\{x_{\partial F}^1 \equiv +1\}} \\ &\quad \left[\int_S \mu(dx^2) 1_{\{x_{\partial F}^2 \equiv +1\}} 1_{B^{z_F}}(x^1 \wedge x^2) 1_{z_\Lambda}(x^1 \wedge x^2)_\Lambda \right], \end{aligned}$$

where $x^1 \wedge x^2$ denotes the configuration

$$(x^1 \wedge x^2)(i) = x^1(i) \wedge x^2(i), \text{ for all } i \in \mathbb{Z}^2.$$

Observe that for fixed x^1 , the mapping $x^2 \mapsto 1_{B^{z_F}}(x^1 \wedge x^2)$ is \mathcal{F}_{F^c} -measurable. Hence we obtain, applying (2.2), for a fixed configuration x^1 ,

$$\begin{aligned} & \int_S \mu(dx^2) 1_{\{x_{\partial F}^2 \equiv +1\}} 1_{B^{z_F}}(x^1 \wedge x^2) 1_{z_\Lambda}(x^1 \wedge x^2)_\Lambda \\ &= \int_S \mu(dy) 1_{\{y_{\partial F} \equiv +1\}} 1_{B^{z_F}}(x^1 \wedge y) \int_{A^\Lambda} p_\Lambda(du_\Lambda|y) 1_{z_\Lambda}(x^1 \wedge u)_\Lambda \\ &= \int_S \mu(dy) 1_{\{y_{\partial F} \equiv +1\}} 1_{B^{z_F}}(x^1 \wedge y) \int_{A^\Lambda} p_\Lambda(du_\Lambda|y_{\partial F}) 1_{z_\Lambda}(x^1 \wedge u)_\Lambda \\ &= \int_S \mu(dy) 1_{\{y_{\partial F} \equiv +1\}} 1_{B^{z_F}}(x^1 \wedge y) \int_{A^\Lambda} p_\Lambda(du_\Lambda|+) 1_{z_\Lambda}(x^1 \wedge u)_\Lambda, \end{aligned}$$

where we used that μ is a Markov random field of order 1 and the fact that $\{j \in \mathbb{Z}^2 : \text{dist}(j, \Lambda) = 1\} = \partial F$, where dist is the distance defined by the L^1 -norm on \mathbb{Z}^2 . Putting things together, we obtain

$$\begin{aligned} \mathbb{P}^\varepsilon(B; X_F = z_F) &= \int_{S \times S} \nu_\varepsilon(dy^1) \mu(dy^2) 1_{\{y_{\partial F}^1 \equiv y_{\partial F}^2 \equiv +1\}} 1_{B^{z_F}}(y^1 \wedge y^2) \cdot \\ &\quad \left[\int_{A^\Lambda} \prod_{j \in \Lambda} [\varepsilon \delta_{-1} + (1 - \varepsilon) \delta_{+1}](dx_j^1) \int_{A^\Lambda} p_\Lambda(dx_\Lambda^2|+) 1_{z_\Lambda}(x^1 \wedge x^2) \right]. \end{aligned}$$

Recall that $z_{\partial F} \equiv +1$. Write for simplicity

$$\varphi_i(\Lambda, z_F) = \int_{A^\Lambda} \prod_{j \in \Lambda} [\varepsilon \delta_{-1} + (1 - \varepsilon) \delta_{+1}](dx_j^1) \int_{A^\Lambda} p_\Lambda(dx_\Lambda^2|+) 1_{z_\Lambda}(x^1 \wedge x^2).$$

Thus,

$$\begin{aligned} \mathbb{P}^\varepsilon(B; X_F = z_F) &= \int_{S \times S} \nu_\varepsilon(dy^1) \mu(dy^2) 1_{\{y_{\partial F}^1 \equiv y_{\partial F}^2 \equiv +1\}} 1_{B^{z_F}}(y^1 \wedge y^2) \varphi_i(\Lambda, z_F) \\ &= \int_S \mathbb{P}^\varepsilon(dy) 1_{\{y_{\partial F} \equiv +1\}} 1_{B^{z_F}}(y) \varphi_i(\Lambda, z_F). \end{aligned}$$

Therefore, if we define

$$p_i^\varepsilon(+1|z) := \frac{\varphi_i(\Lambda, (+1, z_{F \setminus \{i\}}))}{\varphi_i(\Lambda, (+1, z_{F \setminus \{i\}})) + \varphi_i(\Lambda, (-1, z_{F \setminus \{i\}}))}, \quad (3.9)$$

on $\{\bar{C}_i(z) = F\}$, then we have, applying the above arguments to $B \cap \{X_i = +1\}$ with $B \in \mathcal{F}_{\{i\}^c}$, that

$$\int_S \mathbb{P}^\varepsilon(dx) 1_B(x) 1_{z_{F \setminus \{i\}}}(x_{F \setminus \{i\}}) 1_{x_i = +1} = \int_S \mathbb{P}^\varepsilon(dx) 1_B(x) 1_{z_{F \setminus \{i\}}}(x_{F \setminus \{i\}}) p_i^\varepsilon(+1|z).$$

Hence the object defined in (3.9) is a version of the conditional probability $\mathbb{P}^\varepsilon(X_i = +1|z)$ on the event that $\bar{C}_i(z) = F$. It is evident that $z \mapsto p_i^\varepsilon(+1|z)$ is $\mathcal{F}_{C_i(X)}$ -measurable. This concludes the proof.

4 Proof of Theorem 3

This section is devoted to the proof of Theorem 3. The following lemma is the key of our proof.

Lemma 1 *For sufficiently large β and for all ε ,*

$$\mathbb{P}_{\beta,\varepsilon}^+ \left(\bigcap_{i \in \Gamma} \{X_i = -1\} \right) \leq \left(e^{-2\beta} + \varepsilon \right)^{|\Gamma|}$$

holds for any path $\Gamma \subset \mathbb{Z}^2$.

In order to give the proof of the lemma, we need to recall the notion of contour. We first recall the notion of dual.

Definition 9 *We call dual of \mathbb{Z}^2 the set $(\mathbb{Z}^2)' = \mathbb{Z}^2 + (\frac{1}{2}, \frac{1}{2})$. For any finite set $\Lambda \subset \mathbb{Z}^2$, the dual Λ' of Λ is given by*

$$\Lambda' = \{r = (r_1, r_2) \in (\mathbb{Z}^2)' : \exists i = (i_1, i_2) \in \Lambda, r_1 = i_1 \pm \frac{1}{2}, r_2 = i_2 \pm \frac{1}{2}\}.$$

We define the L^1 -norm and the notion of L^1 -neighbors on the dual in exactly the same way as we did for \mathbb{Z}^2 .

A contour is defined as follows.

Definition 10 *Any finite sequence $\gamma = (r_1, \dots, r_n)$ of points in the dual $(\mathbb{Z}^2)'$ is called a contour if $(r_j, \dots, r_n, r_1, \dots, r_{j-2})$ is a path for all $j = 1, \dots, n$. We write $|\gamma| = n$ for the length of the contour. We say that two contours γ and γ' do not intersect if and only if either $R(\gamma) \cap R(\gamma') = \emptyset$ or $\gamma \cap \gamma' = \emptyset$.*

Joining two neighboring points of the contour γ with a straight line, we obtain a closed finite curve in \mathbb{R}^2 . We denote the intersection of its interior with \mathbb{Z}^2 by $R(\gamma)$.

We have now the elements to prove Lemma 1.

Proof of Lemma 1.

To any $(x^1, x^2) \in S^2$, we associate the configuration $x = x^1 \wedge x^2$. Then

$$1_{\{x_i = -1\}} = 1 - 1_{\{x_i^1 = +1\}} + (1 - 1_{\{x_i^2 = +1\}})1_{\{x_i^1 = +1\}}.$$

By definition of $\mathbb{P}_{\beta,\varepsilon}^+$, we have

$$\begin{aligned} \mathbb{P}_{\beta,\varepsilon}^+ \left(\bigcap_{i \in \Gamma} \{X_i = -1\} \right) &= \mu_\beta^+ \otimes \nu_\varepsilon \left(\prod_{i \in \Gamma} \left[1 - 1_{\{x_i^1 = +1\}} + (1 - 1_{\{x_i^2 = +1\}})1_{\{x_i^1 = +1\}} \right] \right) \\ &= \mu_\beta^+ \left(\prod_{i \in \Gamma} [1 - (1 - \varepsilon)1_{\{x_i = +1\}}] \right) \\ &= \sum_{C \subset \Gamma} \mu_\beta^+ (x_C = -1, x_{\Gamma \setminus C} = +1) \varepsilon^{|\Gamma| - |C|}. \quad (4.10) \end{aligned}$$

We use that

$$\mu_\beta^+ = \lim_{\Lambda \rightarrow \mathbb{Z}^2} \mu_{\beta, \Lambda}^+,$$

where $\mu_{\beta, \Lambda}^+ = P_{\beta, \Lambda}(\cdot | +_{\Lambda^c})$, see (2.8), and where $+_{\Lambda^c}$ denotes the configuration $y_j = +1$ for all $j \in \Lambda^c$. In the sequel we will study the properties of the finite volume measure $\mu_{\beta, \Lambda}^+$ and get estimates uniform in Λ , for all Λ containing Γ .

We have

$$\mu_{\beta, \Lambda}^+(x_C = -1, x_{\Gamma \setminus C} = +1) = \frac{1}{Z_{\beta, \Lambda}^+} \sum_{\underline{\gamma} \in \Omega_\Lambda} \prod_{\gamma \in \underline{\gamma}} e^{-2\beta|\gamma|} 1_{\{x_C(\underline{\gamma}) = -1\}} 1_{\{x_{\Gamma \setminus C}(\underline{\gamma}) = +1\}}, \quad (4.11)$$

where

$$\Omega_\Lambda = \{\underline{\gamma} = \{\gamma_1, \dots, \gamma_n\}, n \geq 1, \gamma_i \in \Lambda' \text{ for all } i, \gamma_1, \dots, \gamma_n \text{ non intersecting}\},$$

and where for any given set of contours $\underline{\gamma} \in \Omega_\Lambda$, $x(\underline{\gamma}) \in A^\Lambda$ denotes the associated configuration. In the above formula we used the classical correspondance between configurations and sets of non intersecting contours, see e.g. Presutti (2009).

In order to evaluate (4.11), let $C = C_1 \cup \dots \cup C_n$ be the decomposition of C into the union of its connected components. This means that each C_i is a L^1 -connected set and $\text{dist}(C_i, C_j) \geq 2$ for all $i \neq j$. All components C_1, \dots, C_n must be contained in a contour. Observe that one such contour can contain several components. More precisely, for any m between 1 and n , let $\{J_j, j = 1, \dots, m\}$ be a (disjoint) partition of $\{1, \dots, n\}$ and let

$$P_j = \bigcup_{i \in J_j} C_i.$$

Each of the $P_j, 1 \leq j \leq m$, will be surrounded by exactly one contour γ_j . The contours surrounding different P_j 's have to be non-intersecting. Moreover, the contours $\gamma_1, \dots, \gamma_m$ have to be the only contours that intersect the path Γ .

Given the contours $\gamma_1, \dots, \gamma_m$ as in the last paragraph, write

$$\begin{aligned} \Omega(\Lambda \setminus \gamma_1 \cup \dots \cup \gamma_m) &= \{\underline{\gamma}' = \{\gamma'_1, \dots, \gamma'_k\} \in \Omega_\Lambda : \text{for all } 1 \leq i \leq k, 1 \leq j \leq m, \\ &\quad \gamma'_i \cap \gamma_j = \emptyset \text{ and } R(\gamma'_i) \cap \Gamma = \emptyset\}, \end{aligned}$$

for the set of all contours not intersecting with Γ nor with any of the $\gamma_j, 1 \leq j \leq m$. Now we can rewrite (4.11) as follows.

$$\begin{aligned} &\frac{1}{Z_{\beta, \Lambda}^+} \sum_{\underline{\gamma}} \prod_{\gamma \in \underline{\gamma}} e^{-2\beta|\gamma|} 1_{\{x_C(\underline{\gamma}) = -1\}} 1_{\{x_{\Gamma \setminus C}(\underline{\gamma}) = +1\}} \\ &= \frac{1}{Z_{\beta, \Lambda}^+} \sum_{m=1}^n \sum_{P_1, \dots, P_m} \sum_{\gamma_1: P_1 \subset R(\gamma_1)} e^{-2\beta|\gamma_1|} \dots \sum_{\gamma_m: P_m \subset R(\gamma_m)} e^{-2\beta|\gamma_m|} 1_{\{\gamma_1, \dots, \gamma_m \text{ non-intersecting}\}} \\ &\quad \sum_{\underline{\gamma} \in \Omega(\Lambda \setminus \gamma_1 \cup \dots \cup \gamma_m)} \prod_{\gamma \in \underline{\gamma}} e^{-2\beta|\gamma|}. \end{aligned}$$

Since for any fixed set of $\gamma_1, \dots, \gamma_m$,

$$Z_{\beta, \Lambda}^+ \geq \sum_{\underline{\gamma} \in \Omega(\Lambda \setminus \gamma_1 \cup \dots \cup \gamma_m)} \prod_{\gamma \in \underline{\gamma}} e^{-2\beta|\gamma|},$$

we have

$$\begin{aligned} \mu_{\beta, \Lambda}^+ (x_C = -1, x_{\Gamma \setminus C} = +1) \\ \leq \sum_{m=1}^n \sum_{P_1, \dots, P_m} \sum_{\gamma_1: P_1 \subset R(\gamma_1)} e^{-2\beta|\gamma_1|} \dots \sum_{\gamma_m: P_m \subset R(\gamma_m)} e^{-2\beta|\gamma_m|} 1_{\{\gamma_1, \dots, \gamma_m \text{ non-intersecting}\}}. \end{aligned}$$

Observe that

$$\sum_{i=1}^m \min_{\gamma_i: P_i \subset R(\gamma_i)} |\gamma_i| \geq 2 \sum_{i=1}^m |P_i| + 2n = 2|C| + 2n.$$

As a consequence,

$$\begin{aligned} \sum_{P_1, \dots, P_m} \sum_{\gamma_1: P_1 \subset R(\gamma_1)} e^{-2\beta|\gamma_1|} \dots \sum_{\gamma_m: P_m \subset R(\gamma_m)} e^{-2\beta|\gamma_m|} 1_{\{\gamma_1, \dots, \gamma_m \text{ non-intersecting}\}} \\ \leq e^{-2\beta n} e^{-2\beta|C|} \sum_{P_1, \dots, P_m} \sum_{\gamma_1: P_1 \subset R(\gamma_1)} e^{-\beta|\gamma_1|} \dots \sum_{\gamma_m: P_m \subset R(\gamma_m)} e^{-\beta|\gamma_m|} 1_{\{\gamma_1, \dots, \gamma_m \text{ non-intersecting}\}}. \end{aligned}$$

To obtain an upper bound of the right hand side of the above inequality, we use that

$$\begin{aligned} \bigcup_{P_1, \dots, P_m} \left\{ \{\gamma_1, \dots, \gamma_m\} : P_1 \subset R(\gamma_1), \dots, P_m \subset R(\gamma_m), \gamma_1, \dots, \gamma_m \text{ non-intersecting} \right\} \\ = \left\{ \{\gamma_1, \dots, \gamma_m\} : l_1 \in R(\gamma_1), \min\{l_i : l_i \notin R(\gamma_1)\} \in R(\gamma_2), \dots, \right. \\ \left. \min\{l_i : l_i \notin R(\gamma_1) \cup \dots \cup R(\gamma_{m-1})\} \in R(\gamma_m), \right. \\ \left. \gamma_1, \dots, \gamma_m \text{ non-intersecting and for all } 1 \leq i \leq n, \exists j : C_i \in R(\gamma_j) \right\}, \quad (4.12) \end{aligned}$$

where for every $i = 1, \dots, n$, l_i is a fixed but otherwise arbitrary element of C_i . Hence

$$\begin{aligned} \sum_{P_1, \dots, P_m} \sum_{\gamma_1: P_1 \subset R(\gamma_1)} e^{-\beta|\gamma_1|} \dots \sum_{\gamma_m: P_m \subset R(\gamma_m)} e^{-\beta|\gamma_m|} 1_{\{\gamma_1, \dots, \gamma_m \text{ non-intersecting}\}} \\ \leq \sum_{\gamma_1: l_1 \in R(\gamma_1)} e^{-\beta|\gamma_1|} \sum_{\gamma_2: \min\{l_i: l_i \notin R(\gamma_1)\} \in R(\gamma_2)} e^{-\beta|\gamma_2|} \dots \\ \sum_{\gamma_m: \min\{l_i: l_i \notin R(\gamma_1) \cup \dots \cup R(\gamma_{m-1})\} \in R(\gamma_m)} e^{-\beta|\gamma_m|}. \quad (4.13) \end{aligned}$$

Note that

$$\sum_{\gamma: l_i \in R(\gamma)} e^{-\beta|\gamma|} = \sum_{\gamma: 0 \in R(\gamma)} e^{-\beta|\gamma|}. \quad (4.14)$$

Hence we can upper bound the right hand side of (4.13) by

$$\begin{aligned} \sum_{\gamma_1: l_1 \in R(\gamma_1)} e^{-\beta|\gamma_1|} \sum_{\gamma_2: \min\{l_i: l_i \notin R(\gamma_1)\} \in R(\gamma_2)} e^{-\beta|\gamma_2|} \dots \\ \sum_{\gamma_m: \min\{l_i: l_i \notin R(\gamma_1) \cup \dots \cup R(\gamma_{m-1})\} \in R(\gamma_m)} e^{-\beta|\gamma_m|} \leq \left(\sum_{\gamma: 0 \in R(\gamma)} e^{-\beta|\gamma|} \right)^m, \end{aligned}$$

where we have applied successively the upper bound (4.14) to the right hand side in (4.13), starting with γ_m .

To conclude the proof of the lemma, we need an upper bound for the sum $\sum_{\gamma:0 \in R(\gamma)} e^{-\beta|\gamma|}$. Recall that the number of closed contours of length l that contain 0 is upper bounded by $4l3^{l-2}$. Hence,

$$\sum_{\gamma:0 \in R(\gamma)} e^{-\beta|\gamma|} \leq \sum_{l \geq 4} 4l3^{l-2} e^{-\beta l}.$$

Moreover, for β sufficiently large, $4l3^{l-2} e^{-\beta l} \leq e^{-\beta l/2}$. Hence

$$\sum_{\gamma:0 \in R(\gamma)} e^{-\beta|\gamma|} \leq \frac{e^{-2\beta}}{1 - e^{-\beta/2}} \leq 1. \quad (4.15)$$

We conclude that

$$\mu_{\beta,\Lambda}^+(x_C = -1, x_{\Gamma \setminus C} = +1) \leq \sum_{m=1}^n e^{-2\beta n} e^{-2\beta|C|} = n e^{-2\beta n} e^{-2\beta|C|} \leq e^{-2\beta|C|}, \quad (4.16)$$

for β sufficiently large. Using (4.10), this yields

$$\mathbb{P}_{\beta,\varepsilon,\Lambda}^+ \left(\bigcap_{i \in \Gamma} \{X_i = -1\} \right) \leq \sum_{C \subset \Gamma} e^{-2\beta|C|} \varepsilon^{|\Gamma| - |C|} = \left(e^{-2\beta} + \varepsilon \right)^{|\Gamma|}.$$

Letting $\Lambda \rightarrow \mathbb{Z}^2$, this concludes the proof of the lemma. •

We are now able to give the proof of Theorem 3.

Proof of Theorem 3

Let Γ be a (self-avoiding) path starting at one of the four L^1 -neighbors of the origin. We call this path open if $X_i = -1$ for all $i \in \Gamma$. If $|C_0(X)| = \infty$, then there exist open paths of all lengths starting at one of the four L^1 -neighbors of the origin. Write $N(n)$ for the number of such open paths of length n . The number of such possible paths can be bounded from above by 43^{n-1} . Thus for any $n \geq 1$, using Lemma 1,

$$\begin{aligned} \mathbb{P}_{\beta,\varepsilon}^+(|C_0(X)| = \infty) &\leq \mathbb{P}_{\beta,\varepsilon}^+(N(n) \geq 1) \\ &\leq \mathbb{E}_{\beta,\varepsilon}^+(N(n)) = \sum_{\Gamma:|\Gamma|=n} \mathbb{P}_{\beta,\varepsilon}^+ \left(\bigcap_{i \in \Gamma} \{X_i = -1\} \right) \\ &\leq 43^{n-1} \left(\varepsilon + e^{-2\beta} \right)^n, \end{aligned}$$

and this converges to 0 as $n \rightarrow \infty$, if $2\beta > \ln 3 + e^{-\beta}$ and $\varepsilon < \frac{1}{3} - e^{-2\beta}$.

Concerning the proof of item 2., observe that

$$\begin{aligned} \mathbb{P}_{\beta,\varepsilon}^-(\exists i : |C_i(X)| = \infty) &\geq \mu_{\beta}^-(\exists i : |C_i(X)| = \infty) \\ &= \mu_{\beta}^-(\exists \text{ an infinite open path}) = 1, \end{aligned}$$

by Russo's classical results (see Proposition 1 of Russo (1979)). This concludes the proof.

5 Proof of Theorem 2

For two probability measures μ and ν on $A^{\mathbb{Z}^2}$, write $\mu \preceq \nu$ if there exists a coupling \bar{Q} having μ as first marginal and ν as second marginal, such that $\bar{Q}(\{(x^1, x^2) \in S^2 : x^1(i) \leq x^1(i) \forall i \in \mathbb{Z}^2\}) = 1$.

Note that

$$\mathbb{P}^\varepsilon(X_i = +1 | X_j, j \neq i) \geq (1 - \varepsilon)\lambda_0^+.$$

Now we can apply a standard coupling argument, see for instance Lemma 1.1 of Liggett et al. (1997), to prove that

$$\nu_{(1-\varepsilon)\lambda_0^+} \preceq \mathbb{P}^\varepsilon \quad \text{and} \quad \mathbb{P}^\varepsilon \preceq \nu_{(1-\varepsilon)(1-\lambda_0^-)}.$$

Therefore,

$$\begin{aligned} & \mathbb{P}^\varepsilon(|C_i(X)| = \infty) \\ &= \mathbb{P}^\varepsilon(\text{there exists an infinite path of } -1 \text{ starting from one of the four neighbors of } i) \\ &\leq \nu_{(1-\varepsilon)\lambda_0^+}(\text{there is an infinite path of } -1 \text{ starting from one of the neighbors of } i), \end{aligned}$$

which equals zero by condition (2.6).

In the same way, under condition (2.7),

$$\mathbb{P}^\varepsilon(|C_i(X)| = \infty) \geq \nu_{(1-\varepsilon)(1-\lambda_0^-)}(|C_i(X)| = \infty) > 0.$$

By Kolmogorov's 0-1-Law applied to the product measure $\nu_{(1-\varepsilon)(1-\lambda_0^-)}$, this implies that

$$\nu_{(1-\varepsilon)(1-\lambda_0^-)}(\exists i : |C_i(X)| = \infty) = 1,$$

and hence

$$\mathbb{P}^\varepsilon(\exists i : |C_i(X)| = \infty) = 1.$$

Acknowledgments

We thank two anonymous referees whose remarks helped us to significantly improve the manuscript. We thank D. Y. Takahashi and R. Fernández for stimulating discussions and bibliographic suggestions. This work is part of USP project MaCLinC, “Mathematics, computation, language and the brain”, USP/COFECUB project “Stochastic systems with interactions of variable range” and CNPq project 476501/2009-1. It was partially supported by CAPES grant AUXPE-PAE-598/2011. A.G. is partially supported by a CNPq fellowship (grant 305447/2008-4). E.L. has been supported by ANR-08-BLAN-0220-01. M.C. and E.L. thank NUMEC, University of Sao Paulo, for hospitality and support.

References

- [1] Braitenberg, V., Schütz, A., Cortex: Statistics and geometry of the neuronal connectivity. Springer, Berlin (1998).

- [2] Cessac, B., Statistics of spike trains in conductance-based neural networks: Rigorous results. arXiv: 1104.3795v2 (2011).
- [3] Collet, P., Leonardi, F., Loss of memory of random functions of Markov chains and Lyapunov exponents. arXiv: 0908.0077 (2009).
- [4] Dereudre, D., Drouilhet, R., Georgii, H.O., Existence of Gibbsian point processes with geometry-dependent interactions. *Probab. Theor. Rel. Fields.* **155**, (2011).
- [5] Dobrushin, R. L., Prescribing a system of random variables by conditional distributions. *Theor. Prob. Appl.* **15**, 458–486 (1970).
- [6] Dobrushin, R. L., Perturbation methods of the theory of Gibbsian fields. In *Ecole d'été de Probabilités de Saint Flour XXIV*, Lecture Notes in Mathematics **1648**, Berlin Heidelberg New York, Springer (1996).
- [7] Georgii, H.O., Gibbs measures and phase transitions. de Gruyter, Berlin (1988).
- [8] Grimmett, G.R., Percolation. Springer, Berlin (1999).
- [9] Liggett, T.M., Schonmann, R.H., Stacey, A.M., Domination by product measures. *Ann. Probab.* **25**, 71–95 (1997).
- [10] Löcherbach, E., Orlandi, V., Neighborhood radius estimation in variable-neighborhood random fields. *Stoch. Proc. Appl.* **121**, 2151–2185 (2011).
- [11] MacLean, J., Watson, B., Aaron, G. and Yuste, R., Internal dynamics determine the cortical response to thalamic stimulation. *Neuron* **48**, 811-823 (2005).
- [12] Presutti, E., Scaling limits in statistical mechanics and microstructures in continuum mechanics. Springer Berlin Heidelberg, Series: Theoretical and Mathematical Physics (2009).
- [13] Rissanen, J., A universal data compression system. *IEEE, Trans. Inform. Theory* **29**, 656-664 (1983).
- [14] Russo, L., The infinite cluster method in the two-dimensional Ising model. *Commun. Math. Phys.* **67**, 251–266 (1979).
- [15] Wang, L., Yu, C., Chen, H., Qin, W., He, Y., Fan, F., Zhang, Y., Wang, M., Li, K., Zang, Y., Woodward, T.S., Zhu, C., Dynamic functional reorganization of the motor execution network after stroke. *Brain* **133**, 1224-1238 (2010).

Marzio Cassandro
Dipartimento di Fisica
Università di Roma La Sapienza
P.le A. Moro
00185 Roma, Italy
e-mail: `cassandro@roma1.infn.it`

Antonio Galves
Instituto de Matemática e Estatística
Universidade de São Paulo
Caixa Postal 66281
05315-970 São Paulo, Brasil
e-mail: `galves@usp.br`

Eva Löcherbach
CNRS UMR 8088
Département de Mathématiques
Université de Cergy-Pontoise
95 000 CERGY-PONTOISE, France
email: `eva.loecherbach@u-cergy.fr`